DOCUMENT RESUME

ED 249 295 UD 023 602

AUTHOR Bernal, Ernest M.
TITLE Intelligence Tests on Trial.
PUB DATE 81
NOTE 9p.; Discussion presented at "Intelligence Tests on Trial: Larry P. and PASE" (J. Sattler, Chair), a symposium held at the meeting of the American Psychological Association (Los Angeles, CA, August 1981).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Culture Fair Tests; *Intelligence Quotient; *Intelligence Tests; *Minority Groups; *Nature Nurture Controversy; *Test Bias

ABSTRACT

In the United States, IQ tests are developed by and for whites. IQ tests and their derivates have been used on minorities not so much for prescriptive intervention purposes as for confirmation of suspiciously different behavior and for placement into special education and out of programs for the gifted, higher education, and advanced occupational positions. Efforts to assess and redress cultural bias in IQ tests have been problematic, at best. This is because retrospective analyses of fixed test content are inferior methodologically to a simultaneous analysis of all items in a previously unselected item pool. More basically, however, it must be asked whether the use of IQ tests, even if they were not biased, offers more benefits than disadvantages to the minority student. Although the answer to this question is probably "no," it seems likely that IQ will continue to be measured and that comparative studies of IQ will continue to use extant tests rather than develop new ones. Therefore, a focus on methodological problems in comparison is necessary, including matching and the analysis of environmental versus genetic variables. In addition, the effects of the test taking ambience as a whole must be examined, including a new study of the effects of test taking skills, examiner ethnicity, power tests, and other factors on minority group test performance and test readiness. It is not possible to defend IQ tests without consideration of the problems outlined above. (GC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED249295

MAY 15 1984

Discussion: Intelligence Tests on Trial

by

Ernest M. Bernal, Ph.D.
Creative Educational Enterprises, Inc.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ernest M. Bernal

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

UD023602

Discussion presented at

J. Sattler (Chair), "Intelligence Tests on Trial: Larry P. and PASE," Symposium presented at the meeting of the American Psychological Association, Los Angeles, August 1981.

In the United States, IQ tests are developed by and for Whites, the belated inclusion of minorities in norms tables notwithstanding. IQ tests and their derivates have been used on minorities not so much for prescriptive-intervention purposes as for objective confirmation of suspiciously different behavior and for placement into special education and out of programs for the gifted, higher education, and advanced positions in the world of work. These applications are predictable enough in a society whose ethos is simultaneously normative and competitive and whose history belies a strong if sometimes unconscious racism. If many minority psychologists and others who resist the use of IQ tests on non-dominant ethnic groups don't believe in the utility of these instruments or their fair applications in the field, it is because they have too often found instances of misplacements into special education classes, too often have seen instances of high IQ scorers whose only contribution to knowledge was the doctoral dissertation and whose principal contribution to humanity was their clinical internship, while many productive minority professionals were denied admission to doctoral programs because selection committees would not review any applicants whose GRE scores fell below a specified level.

So much for history! It is important to say these things to clear the air, as it were, about the skepticism which surrounds the use of IQ and related tests, and to allow us to concentrate of the issues of IQ tests bias, the merits of their continued use, and the research interpretations of IQ test results.

I suspect that all well-made tests must incorporate a certain amount of bias in some sense or at some level if they are to be useful. Were this not true, the results of the culture-fair, culture-free movement of some

years ago would not have produced such disappointing results. IQ tests, however, may be biased in ways which state-of-the-art analyses of previously collected data may not be able to study. Other procedures such as those based on factor analysis, may only indicate that minority groups have less than Whites of whatever factor is being measured by the White test on that limited and highly selected set of items. And correlational studies of IQ and school achievement may only reflect that White tests predict well for White institutions.

The best of the "good" IQ tests--popular individually administered tests, such as the Wechsler series--have recently been revised. Most of the items in these revised tests were originally selected long ago at a time when item bias techniques as we now understand them were not used. Nearly three-fourths of the original WISC items, for example, are found in the WISC-R, although one of the major purposes of this revision was to eliminate biased items (Whitworth, 1978). Many of the studies cited by the "Defense" in this debate today are based on various types of bias studies using data derived from the standardization or application of these tests.

I pose for your consideration the issue of the methodological soundness of such retrospective analyses of fixed test content. I maintain that retrospective studies are methodologically inferior to a simultaneous analysis of all the items in a previously unselected item pool, and that retrospective analysis is likely not to yield the same results.

We need to put all these issues in perspective; we need to face the limitations of our psychometric instruments squarely. In this debate on test bias, for instance, no one has suggested an investigation of what the loss of IQ testing for special educational diagnosis has meant to minorities, whether it has hurt or helped the accuracy of assessment and placement.

Dr. Nadine Lambert's comments suggest that in some locales the proportion of minority placements in special education have not been affected by not using these instruments. There may, however, be other effects of which we know nothing. It would appear on the surface, at least, that the use of IQ tests is not worth their expense. The utility and positive benefits of IQ testing from a minority client's or consumer's point of view also needs to be addressed in an empirical manner. Even if IQ tests were not biased, does their use offer more benefits than disadvantages to the minority student? Somehow some of us on the panel seem to assume that making a case for IQ tests' lack of bias is the same as the case for their continued use in field settings on minority groups who have suffered much abuse at the hands of researchers who would readily attribute mean differences in IQ to differences in the genetic potential for intelligence.

Too often psychologists and the lay public both fall prey to the trap of confusing IQ scores and intelligence. It would be interesting to see how Whites would fare on IQ tests developed from scratch by culturally diverse psychologists for their own ethnic groups after carefully validating these measures in various significant social contexts. Such action, however, would play into the very normative-competitive model I earlier decried, and of course would not be very practical in terms of cost, interest, or time.

Hence what will probably happen is that comparative studies of IQ will continue to use extant tests. In an effort to reduce or eliminate certain sources of test bias in the recent past, test developers have taken pains to ensure that normative samples of White and a few minority groups were selected in ways which control certain important variables which the literature tells us affects test performance, such as SES, sex, region of

the country, and urban-rural residence. Indeed, careful studies, such as a recent one done by Reynolds and Jensen (Note 1), matched Blacks and Whites not only on these but also on full-scale IQ (FSIQ) in order to study differences in patterns of mental ability. It is important to remind ourselves that such matchings, while doubtlessly reducing item bias, do not satisfy the more stringent criteria for matching used by geneticists to study phenotypic differences. Groups of Blacks, Whites, Chicanos, Native Americans, and Asian Americans, all members of which would have the same FSIQ, would still have incredible variations within as well as among themselves. Our ability to match or otherwise statistically control for less tangible but, I suspect, equally efficacious variables has limited our ability to explain obtained ethnic differences on IQ. Such variables include the degree of discrimination in the social environment, the linguistic appropriateness of the test, the subject's test-taking skills, and the opportunity in the environment for acculturation or biculturation. There is much more about the personal and social environments which we need to know before we can presume to account for the remainder of systematic variance between or among groups with such constructs as the heritability index (h^2) or with such phrases as the "underlying biological substrate." It remains to be seen whether h^2 can be significantly diminished through a careful study of the effects of such additional environmental variables on tested performance. It also remains to be seen whether h^2 is constant for all ages and groups. Should it prove to be unstable, then perhaps heredity is not what we're talking about at all.

We must bear in mind that the heritability index is but a hypothetical construct to which we, perhaps by default, assign that part of systematic variance which our research designs, resources, and scientific imaginations

cannot otherwise explain. The problem with h^2 is that it is so static: once we assign differences in IQ to differences in genotypes, the issue for many of us dies. Instead of prompting further psychological research, such findings tend to support the work of political activists who would discredit and diminish social programs for the poor and for minority ethnic groups.

We might say, then, that h^2 is a "cop-out": we use it when we cannot explain our data by referring to any other set of variables. I suggest that sociological, anthropological, linguistic, and context-specific ethnographic studies may help us psychologists to identify the environmental variables which will significantly reduce h^2.

I wish to underscore Matarazzo's (1972) review of the heridity-vs.-environment literature on IQ. He found noteworthy the fact that two researchers of different persuasions could examine the same data and reach contradictory conclusions on the relative importance of heredity and environment to measured IQ performance. The heredity-vs.-environment paradigm has failed, in my estimation, to give us satisfactory solutions to the question of racial differences in IQ, and will probably continue to disappoint us in the future, however sophisticated our statistical analyses or designs cast in this manner may become. We need insightful alternatives.

First of all we need to test our assumptions more thoroughly about what our tests measure for ethnic populations, to find out whether the "similar factor structures" discussed here earlier are indeed representative of the same constructs, and whether "parallel regression lines" found in other studies of White and Black differences really speak to the same traits. To do so we may need to begin with more diverse item pools than those we usually develop, those which, as Dr. Jensen points out, load heavily on the general intelligence factor (g).

Then, too, we might explore objective means of calibrating culturally diverse manifestations of intelligence for the full spectrum of intelligence levels, not just the adaptive behavior scales appropriate to the lower end of the distribution.

Earlier I suggested that tests may be biased in ways which item bias analyses may not be able to detect. It is time for us to examine not only individual items and tests but also the effects of the testing ambience as a whole, to study in new ways the effects of test-taking skills and other factors in the literature (such as examiner ethnicity, power tests, etc.) on minority group test performance and on minority group readiness to take the tests on which Whites seem to display such a comperative advantage (Bernal, Note 2). When we have studied these issues we might find that the differences make very little difference after all.

I will conclude by saying that the defense of IQ tests presented by two of the participants is not sufficient to reauthorize their use on minority populations. The relative benefits which should in justice accrue to minority groups by the continued use of these instruments have not been demonstrated. The assumption underlying their arguments seems to be that an unbiased test is reason enough for its widespread application.

On the other hand there appears to be some evidence that IQ tests--for all their supposed lack of bias--have wrought some harm to many minority persons and groups. To re-establish these testing practices without first ensuring against these untoward effects is to risk "burning" yet other generations of minority persons--students, mainly--for the sake of psychometric advancements in realm of IQ.

We might best dedicate our energies in the near future to the creative resolution of these problems, rather than to the defense of IQ tests.

Reference Notes

1. Reynolds, C. T., & Jensen, A. R. Patterns of intellectual abilities among Blacks and Whites matched in "g." Paper presented to the annual meeting of the American Psychological Association, Montreal, 1980.

2. Bernal, E. M. Bias in mental testing: Evidence for an alternative to the heredity-environment controversy. Manuscript submitted for publication, 1981.

References

Matarazzo, J. D. Wechsler's measurement and appraisal of adult intelligence (5th ed.). Baltimore, MD: Williams & Wilkins, 1972.

Whitworth, R. H. Review of WISC-R. In O. K. Buros (Ed.), Eighth mental measurements yearbook (Vol. 1). Highland Park, NJ: Gryphon Press, 1978.